R E P O R T R E S U M E S

ED 019 640 AL 000 834

SUBJECTIVE APPROXIMATION OF RELATIVE LETTER INCIDENCE IN PLEASANT AND UNPLEASANT ENGLISH WORDS.

BY- ANISFELD, MOSHE

EDRS PRICE MF-$0.25 HC-$0.88 20P.

DESCRIPTORS- *PSYCHOLINGUISTICS, COMPUTERS, ENGLISH, PSYCHOLOGICAL TESTING, PHONEMICS, TACHISTOSCOPES,

A COMPUTER COUNT OF THE RELATIVE INCIDENCE OF LETTERS IN 431 PLEASANT (P) WORDS AND 702 UNPLEASANT (U) WORDS REVEALED THAT SOME LETTERS TEND TO OCCUR MORE FREQUENTLY IN THE INITIAL POSITION OF P WORDS AND OTHER LETTERS MORE FREQUENTLY IN THE INITIAL POSITION OF U WORDS. A TASK REQUIRING 40 ADULTS TO GUESS FOR EACH LETTER WHETHER IT OCCURRED IN A P WORD OR IN A U WORD SHOWED THAT PEOPLE ARE ABLE TO APPROXIMATE THESE OBJECTIVE PROBABILITIES OF INITIAL-LETTER OCCURRENCE. THE AUTHOR FEELS THAT THESE FINDINGS CAN EXPLAIN HOW IT IS POSSIBLE TO IDENTIFY THE PROBABLE AFFECTIVE MEANING OF A WORD SEEN IN A TACHISTOSCOPE PRIOR TO ITS COMPLETE RECOGNITION. FASTER RECOGNITION OF P WORDS IN TACHISTOSCOPIC EXPERIMENTS WAS ACCOUNTED FOR IN TERMS OF RESPONSE PROBABILITY. (AUTHOR/DO)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED019640

Subjective Approximation of Relative Letter Incidence in Pleasant and Unpleasant English Words

Moshe Anisfeld

Cornell University, Ithaca, N. Y.

Abstract

A count of the relative incidence of letters in 431 pleasant (P) words and 702 unpleasant (U) words revealed that some letters tend to occur more frequently in the initial position of P words and other letters more frequently in the initial position of U words. A task requiring Ss to guess for each letter whether it occurred in a P word or in a U word showed that people are able to approximate these objective probabilities of initial-letter occurrence. These findings can explain how it is possible to identify the probable affective meaning of a word seen in a tachistoscope prior to its complete recognition. Faster recognition of P words in tachistoscopic experiments was accounted for in terms of response probability.

Running title: Letters in Pleasant and Unpleasant Words

Address: Moshe Anisfeld
Psychology Department
Cornell University
Ithaca, N. Y. 14850

AL 000 834

Subjective Approximation of Relative Letter Incidence in Pleasant and Unpleasant English Words[1]

Moshe Anisfeld

Cornell University

Recent studies (Johnson, Thomson, and Frincke, 1960; Johnson, Frincke, and Martin, 1961; Newbigging, 1961) comparing the tachistoscopic recognition of pleasant (P) and unpleasant (U) words, equated on frequency of usage, support earlier findings of higher thresholds for U than for P words. Such findings have been taken to indicate (e.g., Osgood, 1957) that early in the perceptual process some preliminary identification of the affective nature of the stimulus occurs on the basis of some partial cues and as a result the complete recognition of positive stimuli is speeded up and that of negative stimuli delayed. Eriksen (1960, p. 282) rejected this interpretation on the grounds that no partial cues could convey affective meaning. However, a study by Thorndike (1945) has shown that P and U words differ with respect to at least one part-word characteristic, namely the frequency of occurrence of certain phonemes. Thorndike's study included 6337 phoneme tokens occurring in English words which he evaluated as referring to "...things, qualities, events, or conditions satisfying to the majority of modern civilized men" (P. 143), and 6498 phoneme tokens occurring in English words that refer to "...things and qualities, events or conditions annoying to most civilized men" (P. 143). He found that out of the 43 phoneme types into which he grouped the phoneme tokens, five were

especially frequent in P words and five in U words. Because of lack of one-to-one correspondences in English between letters and phonemes, it is not possible to apply these results in toto to visual perception of words, but an example can be given using the phonemes /b/ and /v/, which are quite regularly represented by the letters B and V, to show how Thorndike's findings are relevant to our concern. Thorndike found that /b/ occurred more frequently in U words and /v/ more frequently in P words. Suppose now that the S in a tachistoscopic experiment detects a V in one word and a B in another word. If this were all the information available to him and he wanted to make intelligent guesses concerning the affective tone of the words, the S's chances of being correct would be greater if he guessed the first word to be P and the second U than if he reversed his guesses. Thus it may not be necessary to recognize the whole word in order to obtain information about its probable affective connotations.

In order to invoke differential probabilities of letter occurrence as an explanation of subthreshold discriminations of the affective character of words, it is necessary to show: (a) that such differential probabilities can be found in the English language--a suggestion derived from Thorndike's results, and (b) that Ss are able to utilize them in appropriate situations. On the latter point the available evidence supports the prediction that Ss would be capable of approximating the differential probabilities of letters in P and U words, if such differences did indeed exist in the language. Howes (1954) reported that Ss could estimate quite accurately the actual frequencies of English words as represented by the Thorndike-Lorge count. Similar results were obtained for bigrams and

trigrams (Mayzner and Tresselt, 1962; Underwood and Schulz, 1960, pp. 52-55) and for single letters (Attneave, 1953).

The purpose of the present study was thus to ascertain if P and U English words were differentiated by selective incidence of letters, and if Ss could draw on such differences when the need arose. To this end, Ss were engaged in two subjective judgment tasks. One was a guessing task in which they were told that letters were drawn from P and U words and they were requested to guess for each letter of the alphabet whether it was drawn from a P word or from a U word. In the second task, Ss were required to estimate directly the relative frequency of each letter in P and U words. The results of these subjective judgments were compared with the findings of an objective count of the relative incidence of each letter in P and U words.

Method

Subjects

Sixty paid volunteers participated in the study. They ranged in age from 19 to 31 years. Most of the Ss were Harvard and Radcliffe students. Of these Ss, 40 (29 males and 11 females) took part in the Guessing and Estimation tasks and 20 (12 males and 8 females) helped in the selection of P and U words for the Count.

Procedure

Word selection for Count. Care was taken to include in the Count only words that were clearly P or U. To achieve this, each word (from preliminary lists to be described) was judged by 10 Ss working individually

as either (a) definitely P, (b) definitely U, or (c) neither definitely P nor definitely U. Words considered by eight or more Ss to fall into categories (a) and (b) were accepted as P and U words respectively. All the other words were not used for the Count. Ss were instructed to sort the words according to the connotations they have for the majority of people and to avoid using idiosyncratic meanings. They were urged to rely largely on their first impressions and to work fairly rapidly.

Two groups of 10 Ss each were employed for this task. The first 10 Ss were given 1150 words for classification. This list of words contained P and U words from published sources and their synonyms obtained from a thesaurus. It consisted of 450 words thought to have pleasant meanings and 700--unpleasant meanings. Each word was typed on a card and the cards were thoroughly shuffled before they were handed to the S. The judgments of this group yielded 280 P words and 605 U words.

Another sample of 340 words was screened by a second group of 10 Ss. These words were obtained from: (a) volunteers who gave lists of 20 P and 20 U words, (b) psychological articles, (c) the Thorndike-Lorge (1944) list of the 1000 most frequent words, and (d) 23 words rated as neither P nor U by the previous group. The judgments of this group resulted in 151 P words and 97 U words. Of the 23 words rated as neutral by the first group of judges, 21 were assigned the same rating by the second group, with 1 P and 1 U. This reflects high reliability.

The combination of the words selected by the two groups yielded a list of 431 P and 702 U word-types. These words were used in one analysis of letter incidence. For another analysis, each word-type was weighted by its frequency in running text as estimated by the Thorndike-Lorge G

tables (A was valued as 50, and AA as 100) to yield P word-tokens and U word-tokens.

Guessing task. On each of 78 index cards, a letter of the alphabet was stencilled, each letter appearing three times. After the 78 cards were thoroughly shuffled, each S was individually shown each letter and requested to guess whether it was selected from a P word or from a U word.

Each S was told that two equally large groups of words had been chosen from a dictionary. One group contained "Pleasant words--words which have pleasant connotations, which bring pleasant thoughts, associations, and memories to mind--for example, the words sweet, unselfish, and successful." The other group contained "Unpleasant words, like sour, selfish, and unsuccessful." (For alternate Ss, the order of introduction of the concepts of "pleasant" and "unpleasant" words was reversed.) The instructions further explained that one letter was picked randomly from each word and S's task was to guess whether the letter belonged to a P or to a U word. It was attempted to make the task meaningful by telling Ss that they 'obviously couldn't be expected to get all the words but there were indications that people could guess correctly a fair number of words in this manner.' They were asked to give their uncensored first impressions.

Frequency estimation task. The second part of the experiment followed the first part directly. In this part, Ss were asked to estimate, on a five-point scale, for each letter of the alphabet whether it occurred with greater or lesser frequency in P or U words. Ss encountered the letters in different random orders.

The order of the two tasks was not counterbalanced as our intention was to base the study primarily on the more sensitive measure of the two.

Results

Letter Frequency Count

The Count was carried out by a 7094 IBM computer. Two analyses were performed: in one, the number of times each letter occurred in P words and U words was determined regardless of its position in the word (Regardless analysis); the second analysis determined the number of times each letter occurred in the initial position in P and U words (Initial analysis). The latter was done because interviews with the Ss after their completion of the Guessing and Estimation tasks indicated that they were heavily influenced by the relative frequencies of initial letters in the P and U categories. Table 1 presents the incidence of each letter in proportion to the total number of letters in the P word-types and in the U word-types and the statistical significance of the difference between the two proportions using the z statistic (Edwards, 1960, pp. 51-54) for both the Initial analysis and the Regardless analysis. Table 2 presents the same information for word-tokens, except that no significance tests were possible in this case because the different tokens of one type are not independent of each other. It can be seen in the Initial analysis column of Table 1 that the letters E and G are significantly more frequent in P word-types than in U word-types and that the letters D, I, M, and O are more frequent in U word-types. The Regardless analysis reveals a significant preponderance of letter E in P word-types and of letters D, M, and S in U word-types.

Insert Tables 1 and 2 about here

In addition to the assessment of the statistical significance of the association of individual letters with P and U word-types, an overall comparison was performed including all letters except X and Z for which there were very few entries, to ascertain the generality of the tendency of letters to occur more frequently in one category of words as against the other. A 24 x 2 chi square analysis comparing the obtained distribution of letters in P and U words with chance expectations yielded a value of 76.51 for the Initial analysis which for df = 23 is significant beyond the .001 level. The chi square for the Regardless analysis is 33.78 (NS). Since P words tend to be more frequent than U words, it was necessary to determine whether the difference in Initial letter distribution was attributable to the P-U variable or to frequency. For this purpose, the A-AA words (138 P words, 66 U words) were analyzed separately from the words with frequencies below 50 per million (292 P, 636 U). The chi square for the Low frequency words was 69.53 ($p < .001$) and for the High frequency words 24.32 (NS). A more balanced division of words into those having frequencies of 16 or higher (264 P, 218 U) and those having frequencies of 15 and under (166 P, 484 U) produced essentially the same results. The chi square for the High category was 29.38 (NS) and for the Low category 58.61 ($p < .001$).

Whatever the explanation for the difference between the High and Low categories, it is important to note for the purposes of the present study, that these two categories were found, in the following analysis, to be drawn from the same population with respect to the P-U variable. The letters in the High category (16+) were classified into the 12 whose

deviations from chance expectations were relatively in favor of the P words (P letters) and the 12 whose deviations were relatively in favor of the U words (U letters), and the same was done for the Low category (15-). Thus a comparison was possible between the letters that fell into the High P and High U categories and those that fell into the Low P and Low U categories. Except for four letters the High and Low groupings coincided, producing a chi-square value of 10.67 ($\underline{df}$ = 1, $\underline{p} < .01$). It can therefore be concluded that the distribution of letters in the initial position of High frequency P and U words is not significantly different from that in Low frequency P and U words.

The conclusion that the differences in letter distribution at the beginnings of P and U words are not mediated through frequency is supported by the lack of correspondence between the 12 P and 12 U letters (determined as above but for the complete sample) on one hand and the 12 most frequent and 12 least frequent letters in initial position (based on the tables of Mayzner and Tresselt, mimeo) on the other hand (chi square = .67, $\underline{df}$ = 1, NS).

The ratio of U/P word-types in our sample is 1.63 (702/403) and of U/P word-tokens .75 (11439/15307) reflecting a greater frequency of usage of the P words. This finding is substantiated by a correlation coefficient of .332 between P (assigned a 2)--U (assigned a 1) and frequency, which for a sample of 1133 words is highly significant, and in accord with previous reports (Goodstein, 1954; Johnson, Thomson, and Frincke, 1960; Koen, 1962).

The correlation between number of letters in a word and P-U is only .047 (mean length: P words = 6.7 letters, U words = 6.5), but it was suspected that this value reflected the counteracting effect of frequency, since frequency and length correlated at -.326 (consistent with Zipf's law).

A partial correlation between P-U and length, removing the influence of frequency, was therefore calculated. The correlation coefficient obtained (.174) is highly significant ($t = 20.28$, $p < .001$). The relative shortness of U words could reflect on adaptive response to a need for a faster verbal reaction to unpleasant stimuli as compared with pleasant ones (for related findings and interpretation, see Pertschonok, 1958).

Guessing and Estimation Tasks

The Ss' responses in the Guessing task were scored in terms of the number of times each letter was attributed to a P word. Since each letter was represented three times in this task, the maximum score an S could get for a letter was 3, i.e., when he placed all three instances of a letter in the P pile, and the minimum score was 0, when he put all three in the U pile. Table 3 displays the values for each letter averaged over the 40 Ss. These values range from 0.73 for the letter U to 2.33 for L. This means that the average S when shown the letter L would almost invariably say it belonged to a P word, and when shown letter U would as certainly assign it to a U word.

Insert Table 3 about here

In the Estimation task the scores ranged from 5, for "much more frequent in P words than in U words," to 1, for "much more frequent in U words than in P words." The values for each letter averaged over all Ss are given in Table 3. As in the Guessing task, the highest mean (3.73) goes to L and the lowest (2.23) to U.

The rank order correlation between the mean Guessing values and the mean Estimation values is .86. Because of this high correlation and a closer relation between the Guessing task and the Count than between the Estimation task and the Count in preliminary analyses, only the former relation will be discussed below.

It was of interest to check whether Ss exhibited in the Guessing task any tendency to favor one category over the other. Since Ss were explicitly told that the letters represented an equal number of P and U words, more letter attributions to one category than to the other would clearly reflect some subjective predisposition. In the absence of bias in either direction, Ss should on the average attribute 36 of the 72 letter occurrences (24 x 3) to P words. The obtained mean for P Guessing is 40.98 and it departs significantly from 36 (t = 6.01, $p < .001$). This finding is not surprising in view of the higher frequency of P words and the correlation between frequency and response probability (Underwood and Schulz, 1960, Ch. 6).

Relation between Performance on Guessing Task and Findings of Count

In order to ascertain whether the Ss' Guesses drew on the actual differences in letter incidence in P and U words as revealed in the four analyses of the Count, it was necessary to determine if Ss tended to give more P responses to letters occurring more frequently in P words and more U responses to those occurring more frequently in U words. For this purpose, the 24 letters (excluding X and Z) were rank ordered on the basis of the differences in their porportions in P and U words, with the rank 1 assigned

to the letter having the highest difference in favor of the P words and 24 to the letter having the highest difference in favor of the U words. The letters were then divided into two equal groups: those having the lowest 12 ranks (P letters) and those having the highest 12 ranks (U letters). The Guessing scores of each S for the 12 P letters were added up as were his scores for the 12 U letters. If the Ss' Guesses were unrelated to the objective differences in letter incidence, the difference between the two sums averaged over all 40 Ss should not depart significantly from zero. The obtained mean difference for Initial position in both the Type and Token analyses is 3.40 (t = 4.61, $p < .001$). (The 12 Initial P letters and the 12 Initial U letters were the same for the Tokens as for the Types.) The mean difference for the Type-Regardless analysis is .40 (t = .36, NS) and for the Token-Regardless analysis 1.48 (t = 1.62, NS).

A similar calculation of the correspondence between the Guessing scores and initial-letter frequency, yielded a mean difference of 4.10 between the Guessing scores for the 12 High frequency letters and the Guessing scores for the 12 Low frequency letters (t = 4.31, $p < .001$). Since no association was found between frequency of letters in initial position and their occurrence in the beginning of P and U words, the two relationships--Guessing with letter frequency, and Guessing with letter distribution in P and U words--would seem to be independent of each other.

Discussion

Tachistoscopic Perception of P and U Words

These results seem to warrant two conclusions: (a) that P and U English words are distinguished by a differential incidence of initial

letters, and (b) that Ss are able to utilize these differences. Applying these findings to tachistoscopic perception, it appears safe to conclude that cues are available in the letter composition of words to allow Ss to infer the probable affective nature of words exposed below threshold for complete recognition. Since in a tachistoscope Ss tend to pick up the beginning of a word first (Bruner and O'Dowd, 1958), they might be expected to reflect, in their anticipation of the affective quality of a word, at least the degree of accuracy indicated by the relation between the Guessing scores and the Initial counts. The level of accuracy could be expected to increase when S isolates two or more letters and their positions, as he might very well before he fully identifies the word.

The present findings provide a simple explanation for the report (Eriksen, Azuma, and Hicks, 1959) that in a tachistoscopic experiment, Ss were able to identify the affective quality of words at a level better than chance even when they could not say what the words actually were. However, sufficient information is not available in the published report of this study to check the validity of this explanation.

It thus appears that partial cues are available in P and U words for a gross prediction of their affective quality. The availability of such cues invalidates the argument as to how a perceiver can reject a stimulus before identifying it, levelled against perceptual defense. However, although now more tenable, the perceptual defense hypothesis is not needed to explain the lower threshold for P than for U words, because this finding can be explained, with the help of results from the present study, by reference to established principles. In experiments

devoted exclusively to the presentation of P and U words, as were the experiments that found faster recognition of P words, S soon learns to expect only words from these categories. These expectations are at least as likely to be biased in favor of P words as were the Guesses in the present study where, despite the emphasis in the instructions on an equal representation of P and U words, Ss attributed significantly more letters to P words than to U words. In a disproportionate number of cases therefore S will try to find a P word that fits the partial stimulus-information perceived. In experiments where the number of P and U words is equal, a portion of the P hypotheses would be erroneously applied to U words and thus lead to a slower recognition of these words. The matching of individual P and U words on frequency does not correct this imbalance, since P words as a class retain their higher frequency and hence higher response probability. The present explanation of the differential thresholds for P and U words rests on notions formulated in the framework of the "new look" in perception (Bruner, 1957; Postman, 1951) and does not require any additional principles.

References

Attneave, F. Psychological probability as a function of experienced frequency. J. exp. Psychol., 1953, 46, 81-86.

Bruner, J. S. On perceptual readiness. Psychol. Rev., 1957, 64, 123-152.

Bruner, J. S. & O'Dowd, D. A note on the informativeness of parts of words. Language and Speech, 1958, 1, pt. 2, 98-101.

Edwards, A. L. Experimental design in psychological research. New York: Holt, Rinehart & Winston, 1960.

Eriksen, C. W. Discrimination and learning without awareness: a methodological survey and evaluation. Psychol. Rev., 1960, 67, 279-300.

Eriksen, C. W., Azuma, H., & Hicks, R. B. Verbal discrimination of pleasant and unpleasant stimuli prior to specific identification. J. abnorm. soc. Psychol., 1959, 59, 114-119.

Goodstein, L. D. Affective tone and visual recognition thresholds. J. abnorm. soc. Psychol., 1954, 49, 443-444.

Howes, D. On the interpretation of word frequency as a variable affecting speed of recognition. J. exp. Psychol., 1954, 48, 106-112.

Johnson, R. C., Frincke, G., & Martin, Lea. Meaningfulness, frequency, and affective character of words as related to visual duration threshold. Canad. J. Psychol., 1961, 15, 199-204.

Johnson, R. C., Thomson, C. W., & Frincke, G. Word values, word frequency, and visual duration thresholds. Psychol. Rev., 1960, 67, 332-342.

Koen, F. Polarization, m, and emotionality in words. J. verb. Learn. verb. Behav., 1962, 1, 183-187.

Mayzner, M. S. & Tresselt, Margaret E. The ranking of letter pairs and single letters to match digram and single-letter frequency counts. J. verb. Learn. verb. Behav., 1962, 1, 203-207.

Mayzner, M. S. & Tresselt, Margaret E. Tables of digram and single letter frequency counts by word length and letter position for a sample of 20,000 words. (mimeo)

Newbigging, P. L. The perceptual redintegration of words which differ in connotative meaning. Canad. J. Psychol., 1961, 15, 133-142.

Osgood, C. E. Motivational dynamics in language behavior. In M. R. Jones (Ed.), Nebraska symposium on motivation. Lincoln: Univer. Nebraska Press, 1957. Pp. 348-424.

Pertschonok, D. The language of feeling: a study in the categorization of emotional expression. Unpublished Ph.D. thesis, Harvard University, 1958.

Postman, L. Toward a general theory of cognition. In J. H. Rohrer, and M. Sherif (Eds.), Social psychology at crossroads. New York: Harper, 1951. Pp. 242-272.

Thorndike, E. L. The association of certain sounds with pleasant and unpleasant meanings. Psychol. Rev., 1945, 52, 143-149.

Thorndike, E. L. & Lorge, I. The teacher's word book of 30,000 words. New York: Teachers College, Columbia Univer., 1944.

Underwood, B. J. & Schulz, R. W. Meaningfulness and verbal learning. New York: Lippincott, 1960.

Footnotes

[1]The assistance of W. Griffin of Harvard University in the collection of data is gratefully acknowledged. Gratitude is also due to R. B. Darlington of Cornell University for his advice on statistical matters. This study was supported by the Harvard University Milton Fund. The computer analysis was carried out by R. Labrie at the Harvard Computation Center and supported by LRI, Harvard Graduate School of Education through the courtesy of J. B. Carroll. The computer facilities were subsidized by National Science Foundation Grant GP-2723. Additional support was provided by the Cornell Faculty Research Grants Committee.

Table 1

Comparison of Proportion of Letters in Pleasant vs. Unpleasant Word-Types in Initial Position and Regardless of Position

Letter	Initial Position			Regardless of Position		
	Proportion in		z	Proportion in		z
	P words	U words		P words	U words	
A	.072	.057	1.0	.080	.083	-0.4
B	.053	.060	-0.5	.018	.023	-1.5
C	.104	.084	1.2	.047	.044	0.7
D	.049	.111	-3.6	.031	.039	-1.9
E	.077	.026	4.0	.144	.120	3.1
F	.046	.070	-1.6	.019	.022	-0.7
G	.072	.028	3.4	.026	.021	1.4
H	.044	.038	0.5	.024	.025	-0.1
I	.035	.061	-2.0	.077	.078	-0.1
J	.016	.007	1.5	.003	.002	0.6
K	.007	.001	1.5	.005	.008	-1.5
L	.044	.030	1.3	.060	.053	1.4
M	.023	.046	-1.9	.024	.031	-1.9
N	.021	.011	1.3	.061	.062	-0.1
O	.000	.026	-3.4	.055	.059	-0.6
P	.081	.061	1.3	.026	.027	-0.1
Q	.002	.004	-0.5	.002	.002	0.0
R	.049	.043	0.5	.078	.080	-0.2
S	.114	.125	-0.6	.058	.072	-2.4
T	.032	.046	-1.1	.073	.065	1.3
U	.007	.017	-1.4	.041	.044	-0.6
V	.023	.023	0.0	.015	.012	1.0
W	.023	.024	-0.1	.008	.008	0.0
X	.000	.000	--	.003	.003	-0.1
Y	.002	.000	1.3	.020	.017	0.8
Z	.002	.000	1.3	.001	.002	-1.2

Note.--A z value of 1.96 is significant at the .05 level and of 2.58 at the .01 level, two-tailed tests.

Table 2

Comparison of Proportion of Letters in Pleasant vs. Unpleasant Word-Tokens in Initial Position and Regardless of Position

Letter	Initial Position		Regardless of Position	
	Proportion in P words	Proportion in U Words	Proportion in P words	Proportion in U words
A	.056	.058	.076	.077
B	.055	.076	.015	.025
C	.095	.086	.041	.040
D	.055	.111	.032	.043
E	.046	.025	.147	.118
F	.062	.099	.023	.027
G	.081	.021	.029	.021
H	.053	.040	.028	.028
I	.020	.034	.065	.067
J	.013	.005	.004	.002
K	.015	.009	.008	.013
L	.049	.024	.057	.060
M	.024	.029	.024	.026
N	.014	.010	.060	.052
O	.000	.008	.059	.061
P	.096	.058	.027	.025
Q	.003	.009	.001	.002
R	.051	.034	.090	.094
S	.114	.134	.062	.069
T	.041	.049	.069	.060
U	.004	.010	.035	.042
V	.016	.018	.013	.012
W	.031	.054	.011	.013
X	.000	.000	.001	.002
Y	.007	.000	.020	.018
Z	.000	.000	.001	.002

Table 3

Mean Values for Letters in Guessing and Estimation Tasks

Letter	Guessing	Estimation
A	2.28	3.68
B	1.50	3.05
C	1.63	2.98
D	1.30	2.48
E	2.28	3.63
F	1.70	2.80
G	1.55	2.95
H	1.75	3.13
I	1.28	3.03
J	1.85	3.58
K	1.15	2.78
L	2.33	3.73
M	2.28	3.40
N	1.45	2.73
O	1.85	3.40
P	1.85	3.68
Q	1.08	2.80
R	1.95	3.20
S	1.93	3.10
T	1.85	2.88
U	0.73	2.23
V	1.45	2.95
W	1.75	3.00
X	0.98	2.28
Y	2.25	3.28
Z	1.33	2.68